MEDINFO 86, R. Salamon, B. Blum, M. Jørgensen (Editors)
Elsevier Science Publishers B.V. (North-Holland)

A Computer-Directed Patient History: Functional Overview and Initial Experience

Peter J. Haug, Homer R. Warner, Paul D. Clayton,
C. Duwayne Schmidt, James E. Pearl, Robert J. Farney

LDS Hospital, Salt Lake City, UT U.S.A.

We have used the medical decision-making tools in the HELP system to develop a frame-based model of pulmonary medical diagnosis. This diagnostic subsystem drives a computer-directed, patient history. It employs a Bayesian scoring algorithm and two levels of control structure. The computer-directed interview collects a history that is comparatively brief and appears to provide diagnostic power equivalent to a more extensive group of pulmonary symptoms collected by paper questionaire. The historical information collected becomes a part of the clinical data base and can be used to revise the statistics that drive the diagnostic frames.

Introduction

Medical information systems were originally developed to improve documentation of patient oriented clinical data and to increase the availability of this data for patient care. Continuing research into expert system design and efforts to integrate medical decision systems with medical information systems have led to applications that supply useful support in daily clinical decision-making. Currently this support is somewhat limited by the absence from the clinical data base of certain information fundamental to medical care. Of this information the most useful is that patient history that reflects the reason for the current admission, ie. the history of present illness.

We have developed a system that uses a frame-based model of medical diagnosis to drive a computer-directed history of present illness. A Bayesian scoring algorithm allows this system to recognize the most likely diseases and choose questions to ask that will be useful in elaborating these diagnoses. A part of the functioning of this system is the generation of a 1 to 5 member differential diagnostic list based on the history collected. This paper describes the functional attributes of and initial experience with the computer-directed patient history.

The system described is part of HELP, a hospital information system functioning in the LDS Hospital in Salt Lake City. Details of the HELP system are have been published previously[1,2], and will be mentioned only briefly here. The foundation of HELP is a large clinical data base, containing a variety of patient information. In addition, a medical decision tool is provided within the system to assist in the interpretation of patient data. Other processes report the results of these interpretations to the appropriate health care personnel. It is thru modifications in the HELP decision tools that the computer-directed history has been developed.

Methods

The HELP decision system is a frame based system in which each frame contains the logical criteria for a particular medical decision. Frames are authored using an editor which allows flexible definition of frame structure and content. We chose to model one disease with each frame and to use these diagnostic frames as hypotheses for a cognitive model[3] of the patient interview. This model uses a cyclic process of hypothesis generation followed by the collection of data necessary to explore these hypotheses as the core of the question generation algorithm.

To restrict the extent of our frame generation task, we chose twenty-eight pulmonary diseases as our initial test diagnostic set. To allow comparison of the relative likelihoods of the diseases under consideration a Bayesian scoring technique was adopted and a group of physicians estimated sensitivities and specificities for a selection of historical findings for each of the diseases. They also estimated an apriori probability of each of the diseases in an inpatient population. Since these frames were designed to participate in a patient interview conducted prior to the availability of other data, laboratory, xray, and physical exam findings were excluded from the diagnostic modules.

Figure 1 is an simplified example of a frame to drive the collection of the history appropriate for the diagnosis of pneumonia. This frame has a structure typical of the diagnostic frames in this experiment.

The structure of probabilistic HELP frames is characterized by several special purpose slots. The frame consists of 1) a title indicating the decision the frame is to make, 2) a slot labled FINAL EVALUATION intended to receive the likelihood generated by the frame, 3) an initial entry under LOGIC containing an apriori probability for the diagnosis represented, 4) a set of SEARCH slots specifying the questions whose answers are required for the evaluation of this diagnosis, and 5) a group of slots that contain the information necessary to calculate a Bayesian probability based on the answer to specific questions. In addition, a special function, the ASK function, is used to indicate which questions should, if their answers are not known, be asked of the patient. This function is preceded by a slot which contains author definable control logic for presenting a question. In the example, the probability following the evaluation of the available information must be greater than or equal to the apriori probability before the frame will cause additional questions to be directed to the patient.

One further aspect of this model is represented in the diagnostic frames. In order for these disease modules to be processed a method must be provided for new data to trigger each diagnostic hypothesis. The frame author indicates these data elements by selecting them and placing a flag (^) in front of the appropriate slots. In the example, an answer to "Have you had recent chest pain?", or "Have you had a cough with this illness?" will cause the logic in the frame for pneumonia to be processed. The initial set of diagnostic hypotheses is built on the answers to 5 fixed questions after which all further questioning is driven by these and subsequent hypotheses.

A separate process, called the QUERY program, is responsible for actually asking the questions of the patient. It runs concurrently with the frame interpreter. After all activated frames are initially processed, the QUERY program gathers the questions sent to a special buffer by the ASK function and determines which five should be directed to the patient. The answers are then stored in the clinical data base and the frames that sent the questions are automatically reprocessed to determine the effect of the new information on the likelihood of those diseases. Other frames may be processed at this time if the new answers trigger them. Thus the answers to a few initial questions result in a iterative process of diagnostic hypothesis generation, hypothesis specific question selection, and as the questions are answered and the frames are reprocessed, resolution of some of the diagnostic hypotheses and the activation of others.

Control Structures

In addition to the diagnostic, question-generating frames described above, certain elements of control structure were needed within the computer-directed history system in order to manage the flow and direction of the questioning process. Two elements of this control structure are embedded in the QUERY program. The first is a question selection algorithm; the second determines when the questioning process can be terminated.

```
(1)   FRAME 1    =.== PNEUMONIA (HISTORY)

      FINAL EVALUATION:
(2)   A VAL:  M

      SECTOR LOGIC:
(3)   A ARITH:     0.014

(4)   B SEARCH: ^ (A) HAVE YOU HAD RECENT CHEST PAIN?

      C SEARCH:   (A) HAVE YOU HAD A FEVER WITH THIS ILLNESS?

      D SEARCH:   (A) HAVE YOU HAD CHILLS WITH THIS ILLNESS?

      E SEARCH: ^ (A) HAVE YOU HAD A COUGH WITH THIS ILLNESS?

      F SEARCH:   (A) IS YOUR CHEST PAIN INCREASED BY BREATHING DEEPLY?
                  (B) IS YOUR CHEST PAIN INCREASED BY COUGHING?
                        USE ANSWER MAX(A, B)

      G SEARCH:   (A) HAVE YOU BEEN SHORT OF BREATH WITH THIS ILLNESS?

      H SEARCH:   (A) IS YOUR SPUTUM YELLOW, GREEN OR BROWN?

(5)   I PROB:     A,  IF ex: C OR D,   USE val: MAX(C, D)
        ANSWER: (N, Y),   TRUE: (0.15, 0.85),   FALSE: (0.7, 0.3)

      J PROB:     I,  IF ex: E,   USE val: E,   ANSWER: (N, Y)
        TRUE:     (0.1, 0.9),   FALSE: (0.8, 0.2)

      K PROB:     J,  IF ex: F,   USE val: F,   ANSWER: (N, Y)
        TRUE:     (0.71, 0.29),   FALSE: (0.9, 0.1)

      L PROB:     K,  IF ex: G,   USE val: G,   ANSWER: (N, Y)
        TRUE:     (0.56, 0.44),   FALSE: (0.87, 0.13)

      M PROB:     L,  IF ex: H,   USE val: H,   ANSWER: (N, Y)
        TRUE:     (0.35, 0.65),   FALSE: (0.95, 0.05)

(6)   N ARITH:   IF M LT A THEN GOTO FINAL EVALUATION

(7)   O EXIST:   ASK((PATIENT QUESTIONS)C, D, E, F, G, H)
```

The question selection process uses an ad hoc estimated of the likelihood that each question in the question buffer will be answered "yes". It is based on the assumption that we will collect the most satisfactory history by attempting to match the data requirements of the most likely diseases. To perform the selection the QUERY program sums for each question in the question buffer the likelihoods of the disease frames that sent that question to the question buffer. Thus questions whose answers would contribute to more than one diagnosis tend to score higher that questions that are used by a single diagnostic hypothesis. The process then compares the totals for each question and selects the top 5. These questions are directed to the patient.

The history termination algorithm is designed to end questioning when it appears unlikely that any further useful data will be found. It is based on the assumption that, if after a reasonable number of questions are answered the total likelihood of unexplored disease is low, further questioning is unlikely to gather additional useful information. Based on this assumption the QUERY program begins by asking the first 30 questions generated by the above processes. Thereafter, following each group of 5 questions, it sums the probabilities of the unexplored diseases that have questions remaining in the question buffer. If the sum of the probabilities of these hypothesized but unexplored diagnoses is less than 5%, the QUERY program terminates the questioning process.

Thus two levels of control are incorporated into the computer-directed history system. A local level of control is specified by the frame author when he first sets the flags specifying which data will trigger each frame and when he specifies the logic that can interrupt frame processing prior to the ASK function. A global control structure manipulates question flow and direction after the frames send questions to the question buffer.

Bayes

The structure that controls the flow of the computer-directed history could be coupled to a variety of scoring algorithms. Thresholds in the QUERY program are adjustable so that any scoring function could be used to define the relative likelihoods of the diseases modeled. We chose to use a Bayes-based scoring approach for three reasons.

The first is the experience with sequential Bayesian diagnosis both at our institution and elsewhere. DeBombal[4], Zagoria[5], Engle[6], and others have described promising computerized diagnosticians based on Bayesian scoring. In our institution, Warner has described the use of Bayesian probabilities to diagnose congenital heart disease[7] and, in a predecessor to this project, to a driver for a program for collecting a screening history in the multiphasic screening area[8].

Figure 1: Parts of a diagnostic frame for the computer-directed history: *(1)* frame label, *(2)* final evaluation slot, *(3)* apriori probability for this disease, *(4)* data specification; indicates the questions reqired to calculate disease likelihood, *(5)* specification of statistics (sensitivity and specificity) associated with yes and no answer to referenced question, *(6)* control logic for ASK function, *(7)* specification of questions to ask patient.

A second consideration is the potential to improve the behavior of Bayesian scoring algorithms through the analysis of data bases. The major objection to the use of Bayesian scoring algorithms is the lack of accurate statistics. In our system, where the statistics in the initial version were estimated by medical experts, we have the ability to generate more accurate values from the historical data base created through the use of the program itself. Initial analysis of a collection of 500 pulmonary histories suggests that the estimates of the sensitivities and specificities of individual patient-oriented questions in our prototype were off by an average of 10 percentage points. The replacement of the estimated statistics with the calculated statistics gave a improvement in diagnostic accuracy from 79% to 88% for the patients in this data base. We are currently verifying this result with a fresh group of patients.

Another objection to the sequential use of Bayes has been the tendency of medical reality to violate the assumptions underlying Bayes equation. Truely accurate sequential calculation of Bayesian likelihoods is based on the underlying assumptions of mutually exclusive diseases and the conditional independency of the data used. Not only do hospitalized patients frequently have multiple diseases but historical data is often lacks independence.

We have attempted to ameliorate these effects in two ways. First the version of Bayes equation used in this project is designed to avoid the pitfalls of using Bayes to compare diseases that are not mutually exclusive. We have used:

$$P_{D/S} = \frac{P_D * P_{S/D}}{P_D * P_{S/D} + P_{ND} * P_{S/ND}}$$

where $P_{D/S}$ is the probability of disease **D** given symptom **S**, P_D is the apriori probability of disease **D**, $P_{S/D}$ is the sensitivity of symptom **S** for disease **D**, P_{ND} is the probability of not having disease **D**, and $P_{S/ND}$ is the false positive rate (1 - the specificity) for symptom **S** in disease **D**. Use of this version of Bayes results in each disease being scored independently and fits the system better to inpatient medical profiles where multiple diseases are common.

Adjusting for the conditional interdependence of symptoms in human disease is more difficult and is done on an individual basis. In our example of the frame for pneumonia we see, for instance, that two questions designed to elicit pleuritic chest pain, "Is your chest pain increased by breathing deeply?" and "Is your chest pain increased by coughing?" are effectively "or'ed" by including them in the same SEARCH statement. The results of the combined answers to these interdependent questions are calculated using a single Bayesian construct.

Experience

In a test of the effectiveness of the computer-directed history program the frames representing the twenty-eight pulmonary diseases listed in table 1 were used to drive history collection at the bedsides of a group of inpatients. In 36 of these patients the discharge diagnoses ultimately showed 1 or more pulmonary diseases. A second group of patients received a paper questionnaire on which they answered all of the 182 questions that would have been required to diagnoses all 28 diseases. 31 of these patients ultimately had one or more pulmonary diseases listed as discharge diagnoses. Using the data collected, the system was asked to generate a differential diagnostic list consisting of the 0 to 5 most likely diagnoses based on history alone for each patient. Table 2 summarizes the results.

Table 1: Diseases for which diagnostic models were created

Acute Bronchitis	Histiocytosis X
Asbestosis	Hodgkin's Disease
Aspiration Pneumonia	Influenza
Asthma	Lung Abscess
Bacterial Pneumonia	Metastatic Neoplasm
Bronchiectasis	Non-Hodgkin's Lymphoma
Chronic Bronchitis	Primary Pulmonary Neoplasm
Coal Worker's Pneumoconiosis	Primary Pulmonary Hypertension
Coccidioidomycosis	Pulmonary Embolism
Congestive Heart Failure	Sarcoidosis
Diffuse Idiopathic Fibrosis	Silicosis
Drug Related Pneumonitis	Spontaneous Pneumothorax
Emphysema	Tuberculosis
Goodpasture's Syndrome	Wegner's Granulomatosis

The patients who took the computer-directed history had a total of 56 pulmonary diseases. Those who had their history collected by questionnaire had a total of 41 diseases . The computer-directed history reduced the number of questions asked to a mean of 78 ± 30 (mean ± SD) yes/no questions per patient. A disease found in the patients discharge diagnosis was correctly identified in the differential list for those with computer collected histories 71% of the time (40 of 56) and in those with histories collected by questionnaire 52% of the time (21 of 41). (The change in accuracy was nonsignificant: p=.069.)

Since some of these patients had multiple pulmonary diseases, we examined how often each differential diagnostic list was completely accurate. This occurred when all of the discharge pulmonary diagnoses were included in a list. For the patients whose history was gathered by questionnaire 61% (19 of 31) of the computer-generated diagnostic lists were completely accurate; for the computer-directed history 66% (23 of 35) of the lists contained all the diagnoses (p > .50).

In addition, there was a tendency for the computer-directed history to respond to more complex presentations, as indicated by more pulmonary diseases in a single patient, by asking greater numbers of questions. It asked a mean of 40 ± 24 questions of patients without pulmonary disease, 73 ± 36 questions of those with one pulmonary disease, and 85 ±27 questions in cases of multiple pulmonary diseases.

Table 2: Accuracy of Diagnoses with Different Data Collection Modes

	Patients with Questionaire Collected History	Patients with Computer-Directed History
Number of Patients	31	35
Number of Pulmonary Diseases	41	56
Number of Accurate* Diagnoses	21 (52%)	40 (71%)**
Number of Questions Asked	182	78

*Accuracy determined by the presence of the correct diagnosis in the differential diagnostic list (see text).
**p=.069.

Conclusion

Patient history provides a majority of the information used to establish a differential diagnostic list in most clinical settings. This differential diagnostic list in turn gives direction to the remainder of the work-up and therapy. History has been notoriously lacking from most hospital information systems because of difficulties with its routine acquisition. We have attempted to ease this data acquisition problem by designing a system capable of directly questioning inpatients using a "intelligent" model of question selection. This system is based on medical knowledge embedded in a set of disease oriented frames created through an interactive frame editor. It has proven capable of collecting patient histories through a terminal located at the patient bedside.

We have chosen to evaluate the function of this system by comparing the diagnostic lists generated by the frame's logic to the discharged diagnoses determined at the time of patient discharge. The only data made available to these frames was the patient history. Initial evaluation shows no loss of diagnostic power when the system chooses a subset of the possible questions to ask the patients. In fact diagnostic accuracy increased although not significantly.

Since these processes generate a data base of histories for patients whose discharge diagnoses are also capture by the HELP system the ability to increase the accuracy of the hypothesis evaluation process is inherent in the system. Plans are underway to automate this process by developing programs which will examine this data base and update the statistics that drive the Bayesian scoring algorithm. These programs will, in a sense, allow the system to learn from its own history gathering experience.

We are continuing the development of the computer-directed history system. The long term goals of this experiment are three fold. We plan to expand the history system to encompass diagnoses from other areas of medicine, to refine the control logic in an effort to capture the maximum diagnostic information while minimizing the patient's burden of answering questions, and to integrate the standard clinical data collected by the HELP system into the diagnostic frames. We hope to create a tool which will not only contribute to the completeness of the patient data base but will also lay the ground work for continued research into hypothesis-directed collection of medical information.

[1]Warner, H.R.: Computer-Assisted Medical Decision-Making. Academic Press, N.Y, New York (1979)

[2]Pryor, T.A., Gardner, R.M., Clayton, P.D., and Warner, H.R.: The HELP System. Journal of Medical Systems (1983); 7:87-102

[3]Elstein, A.S., Shulman, L.S., and Sprafka, S.A.: "Medical Problem Solving: An Analysis of Clinical Reasoning." Harvard Univ. Press, Cmbridge, Mass., 1978

[4]DeDombal, F.T., Leaper, D.J., Horrocks, J.C., Staniland, J.R., and McCann, A.P.: Human and Computer-aided Diagnosis of Abdominal Pain: Further Report with Emphasis on Performance of Clinicians. British Med. Journ. (March 2, 1974); 1:376-380

[5]Zagoria, R.J., Reggia, J.A., Price, T.R., and Banko, M.: Bayesian Classification in Medicine: The Transferability Question. Proc. Fifth Ann. Symp. on Comp. Applic. in Med., IEEE Press, Nov. 1981

[6]Engle, R.L., Flehinger, B.J., Allen, S., Friedman, R., Lipkin, M., Davis, B.J., and Leveridge, L.L.: Heme: A Computer Aid to Diagnosis of Hematological Disease. Bull. N.Y. Acad. Med. (June 1976); 52:584-600

[7]Warner, H.R., Toronto, A.F., and Veasey, L.G.: Experience with Bayes' Theorem for Computer Diagnosis of Congenital Heart Disease. Ann. N.Y. Acad. Scien. (July 31, 1964); 115: 558-567

[8]Warner,H.R., Rutherford, B.D., and Houtchens, B. : A Sequential Bayesean Approach to History Taking and Diagnosis. Computers and Biomedical Research(Jan, 1972) 5:256-262.